Models

What is a model?
1. A description or analogy used to visualize something
2. A system of postulates, data and inferences presented as a mathematical description of an entity or state of affairs

What are models used for?
1. Synthesize observations
2. Explain observations
3. Make observations
4. Sensitivity analysis (how sensitive is a system to a change in one variable)

Models are simplifications of reality, but provide means to ascertain internal consistency in our attempts to interpret observations and to test the plausibility and compatibility of hypothesis concerning the interactions of key processes in a system.

Types of Models

Conceptual
Often the most useful starting point is to construct a reasonably detailed **qualitative** picture of the processes and interactions that occur in the system of interest

In some disciplines such as geology and geography it is difficult to move past conceptual models because of the complexity of the system in question and lack of data to constrain a quantitative model.

Empirical

The term empirical refers to originating in or based on our observations or experiences, relying on observation or experiences often without due regard to system or process. Thus empirical models are those that describe relationships that are observed. Empirical models are often, though not always statistical in nature. They are quite common and often are used to relate a physical parameter that is hard to measure directly to one that is.

Some things to remember about empirical models:
1. in statistics correlation does not necessarily imply causation. This means that while two things are correlated that they are not necessarily linked in a physical sense.
2. The independent variable (measured variable) and the dependent variable (predicted variable) should be related in a meaningful way. For instance, I am interested in runoff from glaciers in Nepal. I may be able to find a strong correlation between the glacier runoff and the price of tea in Nepal, but this correlation tells me nothing about the system I am interested in so it is a worthless empirical model. A more complex example is also of concern in glaciology. It has long been recognized that glacier runoff is highly positively correlated to air temperature (the higher the air temperature, the greater the runoff). However, physical studies of the melting of glaciers show that solar radiation provides the most energy for the melting of the ice. However, the correlation between net radiation and runoff may be lower than for air temperature. This is simply because air temperatures show higher fluctuations than net radiation and because both air temperature and runoff are related to net radiation, there is a high correlation between them, but physically the relationship between net radiation and runoff is more important.

A simple example of determining the correct 'physical' relationship between variables can come from my own research. A large problem in glaciology is determining the volume of ice in an area, however this is notoriously difficult to measure directly. Thus it is common to try to relate glacier area, which is much easier to measure, to glacier volume. Another student and I took different approaches to the problem of relating glacier volume to area.

The other student took the statistical approach to determine an empirical relationship. Unfortunately this relationship is markedly non-linear. This is to say that the relationship cannot be modeled by a linear equation (e.g. $y = mx + b$). Instead, the student used a relationship of the form ($y = a + bx + cx^2$), which is commonly done in statistics.

What always bothered me was why glacier volume would scale as a function of the area^2. I would think that it would scale more as a function of area3/2, because for a cube with the length of a side L, the area increase as L*L and the volume increases as L*L*L. In fact studies of modern glaciers suggest that volume is related to area as:

$V = cA^{1.36}$

Thus, while the statistical correlation for both relationships (see figure below) are both quite high, there is a physical reason to prefer one empirical model to the other.
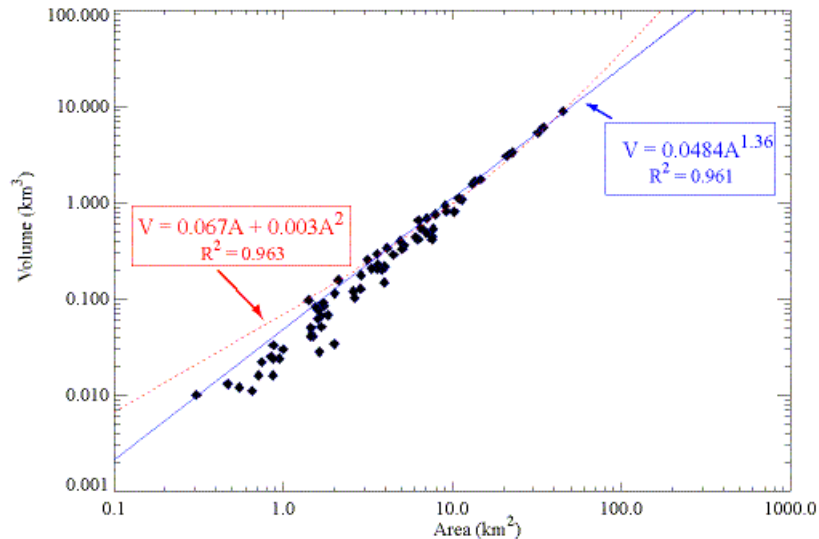
Physical Models

A plethora of physically based models exist. These models are based on known physical (or biological relationships) that provide quantitative relationships. We will only explore one type of physical models- **climate models**. These are important because they are our major tool for predicating the effects of changes in greenhouses gases on the global climate, as wells for more mundane, but important purposes, such as weather forecasting.

Climate Models

A variety of different types of climate models exist. No matter how complex they seem, they are all simplifications of reality. Models are simplifications of reality in two ways. First the physical processes in the model are simplified from those actually occurring. Secondly, the time and space scales of these models are large (e.g. climate models are not run globally at 1-meter spatial resolution every second of the day).

Climate models are designed to simulated today's climate, they can be tested against past climates (e.g. glacial maximum) and are used to predict future climate change.

Important Quantities

Radiation - both solar and terrestrial radiation

Dynamics - *horizontal* movement of energy around the globe by winds and ocean transport from low to high latitudes and *vertical* motions in the atmosphere are considered.

Surface Processes - both land and ocean surfaces are considered

Chemistry

Resolution - both spatial and temporal

Types of Climate Models
1. Energy balance models
    - zero or one-dimensional. A zero dimensional model represents the earth as a single point, while a 1-D model usually predicts variation of temperature with latitude.
    - There are lots of 1-D models and we will be developing one later. They originally showed that small perturbations in the amount of solar radiation will lead to a totally ice-covered earth.

2. 1-D Radiative-Convective Models
    - Compute a vertical profile (usually globally averaged) by explicit modeling of radiative processes and convective adjustment for temperature changes with elevation (lapse rates).
    - the main emphasis of these model is explicit calculation of solar and terrestrial radiation fluxes
    - Usually calculate a global averaged temperature, but also can be used to determine the temperature at various levels in the atmosphere.

3. 2-D Statistical Dynamical Models
    - treat surface processes and dynamics in a zonally-averaged fashion.
    - can be either 2 horizontal or 1 horizontal and 1 vertical dimension
    - use a statistical representation of wind speed and direction and the concept of eddy diffusion to model heat transport
    - largely replaced by GCMs

4. GCMs
    - These are what most people think of as climate models. The acronym variously refers to 'global climate model', 'general circulation model', 'global-circulation model'.

- These models are fully 3-D and may include complex interactions between the atmosphere and ocean and land surface.
- These models solve a series of governing equations at a series of grid cells and at a fixed time step. The equations are non-linear and are solved numerically.
    - spatial resolutions of 4° x 5° or 2° x 2.5° is common today
    - time steps of 15 to 30 minutes are common
    - ~20 vertical levels in the atmosphere are common
- Ocean/Atmosphere dynamics are a major problem and major area of development of these models today - coupled Ocean-Atmosphere Models. The problem is twofold. First, eddies in the ocean are ~10 to 50 km while atmospheric eddies are ~1000 km. Obviously to resolve ocean dynamics a much finer spatial resolution is needed. Also the ocean takes much longer to respond to changes than the atmosphere.
- Hard to represent small scale phenomena (like cloud microphysics or thunderstorms) in GCMs
- Coupling the land surface and atmosphere is difficult as well
- the complexity of the equations make computing power a primary limitation
- A major problem with GCMs is adequate representation of the physics of important processes. Because the physics are so complex or so unknown, sometimes they cannot be adequately represented in GCMs. The process of neglect/semi-empirical or inaccurate representation of processes is called **paramaterization**. In its simplest form, something is simply ignored - *null paramaterization*. Another major form is called *climatological simplification*, which is applying a fixed or non-interactive boundary condition (e.g. fix the temperature of the sea surface to measured values). However for processes that have been parameterized in this manner no feedback mechanism is possible. A different, but still problematic parameterization scheme is to *tune* the parameters in the model to correctly reproduce present day conditions. However, this may lead to unrealistic responses to climatic forcings. The most advanced parameterizations are physically based. It is necessary to make sure that the processes are evenly balanced. For instance, don't include one process in a model and not include another, possibly offsetting processes. An example would be clouds that both reflect energy to space and trap terrestrial radiation.
- Interactions - the relative importance of processes and the manner in which they interlink is strongly dependent on the timescale being modelled. Early on in modeling efforts the timescale of interest is fixed so that processes that are of concern can be identified. Things that vary on timescales shorter than those of

interest can usually be considered adding random variability to the timescale of interest, while those that change on longer timescales than those of interest can usually be considered fixed 'boundary conditions'.

5. Mesoscale Models
    - Similar to GCMs, but with much finer spatial resolution used for modeling and prediction of *regional* climates (e.g. Chesapeake Bay).

6. Data Assimilation Models.
    - These models use atmospheric dynamics to integrate (assimilate) numerous types of climate information into a gridded global data set.